Lieff
Cabraser
Heimann&
Bernstein
Attorneys at Law

SUSMAN GODFREY L.L.P.
A REGISTERED LIMITED LIABILITY PARTNERSHIP

CDAS

October 23, 2024

VIA ECF

Hon. Ona T. Wang
Daniel Patrick Moynihan
United States Courthouse
500 Pearl St.
New York, NY 10007

> RE: *Authors Guild v. OpenAI Inc.*, 23-cv-8292 (S.D.N.Y.) and *Alter v. OpenAI Inc.*,
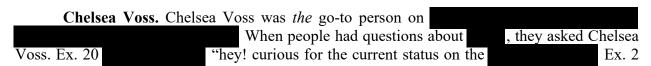> 23-cv-10211 (S.D.N.Y.): Custodians

Dear Judge Wang:

Pursuant to Rule II(b) of Your Honor's Individual Practices, Plaintiffs seek a conference and request that the Court order OpenAI to collect and produce ESI from eight custodians.[1] In September, Plaintiffs moved to compel OpenAI to review and produce the files of six custodians. In its September 13 Order, the Court ordered OpenAI to preserve the files of those custodians. Dkt. 202. After further review, Plaintiffs continue to believe that their files should be produced, and therefore renew their motion. Plaintiffs further request two additional custodians.
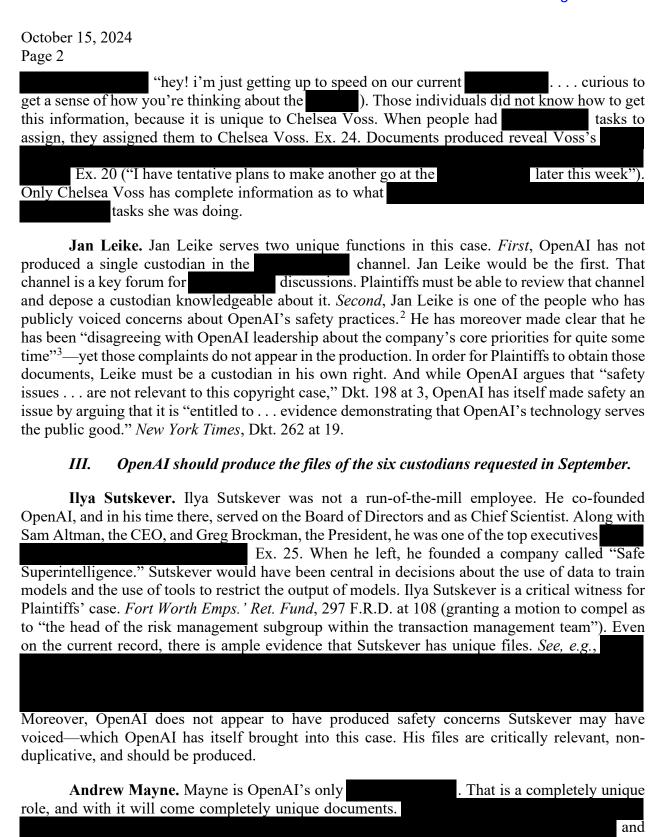
> ### I.   *Overlapping documents do not shield these custodians' files from disclosure.*

To each custodian Plaintiffs have requested, OpenAI's only response is that the custodian is duplicative, because some of their documents also involve other, already-designated custodians. Dkt. 198 at 3. OpenAI does not dispute the relevance of the custodians or substantiate any undue burden. Its reliance on that overlap is misplaced. If there is overlap, the burden on OpenAI is minimal—standard "procedures [exist] to eliminate duplicative search output." *Fort Worth Emps.' Ret. Fund v. J.P. Morgan Chase & Co.*, 297 F.R.D. 99, 106 (S.D.N.Y. 2013); *see Garcia Ramirez v. U.S. Immigr. & Customs Enf't*, 331 F.R.D. 194, 197-98 (D.D.C. 2019). Plaintiffs seek only non-duplicative documents that hit on the agreed search terms. Given that the requested custodians are key decisionmakers and protagonists in OpenAI's conduct and will almost certainly be deposed, their files are highly relevant, critical to taking those depositions, and should be produced.

> ### II.   *OpenAI should produce the files of Chelsea Voss and Jan Leike.*

Chelsea Voss. Chelsea Voss was *the* go-to person on ██████████ ████████████████████████ When people had questions about ██████, they asked Chelsea Voss. Ex. 20 ████████████ "hey! curious for the current status on the ███████████ Ex. 2

---

[1] The parties have met and conferred about the relief sought and did not reach agreement.

October 15, 2024
Page 2

███████████████ "hey! i'm just getting up to speed on our current ███████ . . . . curious to get a sense of how you're thinking about the ███████). Those individuals did not know how to get this information, because it is unique to Chelsea Voss. When people had ███████ tasks to assign, they assigned them to Chelsea Voss. Ex. 24. Documents produced reveal Voss's ██ Ex. 20 ("I have tentative plans to make another go at the ████████████ later this week"). Only Chelsea Voss has complete information as to what ████████████████ ████████████ tasks she was doing.

Jan Leike. Jan Leike serves two unique functions in this case. *First*, OpenAI has not produced a single custodian in the ████████████ channel. Jan Leike would be the first. That channel is a key forum for ████████████ discussions. Plaintiffs must be able to review that channel and depose a custodian knowledgeable about it. *Second*, Jan Leike is one of the people who has publicly voiced concerns about OpenAI's safety practices.[2] He has moreover made clear that he has been "disagreeing with OpenAI leadership about the company's core priorities for quite some time"[3]—yet those complaints do not appear in the production. In order for Plaintiffs to obtain those documents, Leike must be a custodian in his own right. And while OpenAI argues that "safety issues . . . are not relevant to this copyright case," Dkt. 198 at 3, OpenAI has itself made safety an issue by arguing that it is "entitled to . . . evidence demonstrating that OpenAI's technology serves the public good." *New York Times*, Dkt. 262 at 19.

III.   OpenAI should produce the files of the six custodians requested in September.

Ilya Sutskever. Ilya Sutskever was not a run-of-the-mill employee. He co-founded OpenAI, and in his time there, served on the Board of Directors and as Chief Scientist. Along with Sam Altman, the CEO, and Greg Brockman, the President, he was one of the top executives ████████ ████████████████████████████████ Ex. 25. When he left, he founded a company called "Safe Superintelligence." Sutskever would have been central in decisions about the use of data to train models and the use of tools to restrict the output of models. Ilya Sutskever is a critical witness for Plaintiffs' case. *Fort Worth Emps.' Ret. Fund*, 297 F.R.D. at 108 (granting a motion to compel as to "the head of the risk management subgroup within the transaction management team"). Even on the current record, there is ample evidence that Sutskever has unique files. *See, e.g.*, ████ ██ ██ ██ ██

Moreover, OpenAI does not appear to have produced safety concerns Sutskever may have voiced—which OpenAI has itself brought into this case. His files are critically relevant, non-duplicative, and should be produced.

Andrew Mayne. Mayne is OpenAI's only ████████████████ . That is a completely unique role, and with it will come completely unique documents. ████████████████████ ██ and conversations about the above coming from Mayne's unique perspective. Plaintiffs subpoenaed

---

[2] https://x.com/janleike/status/1791498184671605209
[3] https://x.com/janleike/status/1791498178346549382

October 15, 2024
Page 3

Andrew Mayne; he directed Plaintiffs to seek relevant documents from OpenAI. Ex. 26.

Shantanu Jain. As shown in Plaintiffs' underlying motion, ███████████ ████████████████████ OpenAI did not dispute that in its motion. Instead, it argued only that it had "already identified numerous custodians on these subjects," and that "the documents Plaintiffs cite in support of this request include some of those designated custodians." Dkt. 198. They ignore the fact that one of those documents specifically showed that Jain ███ ████████████████████ Other documents likewise show that Jain was having relevant conversations in channels with no designated custodians. ██████████████ ███ ██████████████████████████████████████

Qiming Yuan. ██ ████████████████ Only his files will reflect the full scope of the analyses he conducted. He was the person to whom Voss turned to ██████████████████████████ Ex. 2. Only his files will explain that process. Voss also turned to Yuan with questions relating to the ██████████, Ex. 5, because again, Yuan had unique knowledge in that regard.

Jong Wook Kim. Jong Wook Kim was involved not only in the ████████████ ███ ███████████████████████████ he appears to have been the point person: When a colleague came across a ██████████████████ he sent it to Kim specifically. Ex. 7. Notably, that thread lacks any other designated custodian. ██████████████ ███ Plaintiffs will not be able to capture Kim's documents or relevant ██████████████ information unless he is designated as a custodian.

Cullen O'Keefe. While at OpenAI, O'Keefe did governance work—in his own words, he "worked at OpenAI and worked on policy for GPT-4."[4] In that capacity, he drafted ██ ███ O'Keefe continued to do AI governance work after he left OpenAI.[5] As explained above, OpenAI has made AI governance an issue in this case. Yet they have not named a *single* custodian whose primary role was governance. In opposition to O'Keefe, OpenAI said that they had already named three custodians mentioned in the ██████████████████ Dkt. 198. That is beside the point. The fact that O'Keefe sourced factual information from relevant custodians does not change the fact that his own role was unique among the custodians identified, and that he will have relevant documents in that capacity. OpenAI also argues that O'Keefe was, at one point, a lawyer, and so many of his documents will be privileged. But it is undisputed that that was not his only role at the company, and OpenAI has refused to identify the dates during which O'Keefe served in a legal capacity, either during the parties' conferrals or thereafter. OpenAI cannot rely on privilege to reject O'Keefe wholesale.

---

[4] https://juralnetworks.substack.com/p/ai-safety-is-sometimes-a-model-property
[5] https://cullenokeefe.com/

October 15, 2024
Page 4

For the foregoing reasons, OpenAI should search and produce these custodians' files.

Sincerely,

LIEFF CABRASER HEIMANN
& BERNSTEINS LLP

SUSMAN GODFREY LLP

COWAN, DEBAETS,
ABRAHAMS & SHEPPARD
LLP

/s/ Rachel Geman
Rachel Geman

/s/ Rohit Nath
Rohit Nath

/s/ Scott J. Sholder
Scott J. Sholder